

СРАВНЕНИЕ АЛГОРИТМОВ СТАТИСТИЧЕСКОЙ ОБРАБОТКИ ДАННЫХ МЕЖЛАБОРАТОРНЫХ ЭКСПЕРИМЕНТОВ, ИЗЛОЖЕННЫХ В ГОСУДАРСТВЕННЫХ СТАНДАРТАХ ГОСТ 8.532-85 И ГОСТ 8.532-2002

Е.А.Анчутина, Л.Л.Петров
Институт геохимии им. А.П.Виноградова СО РАН
Иркутск, Фаворского 1а
anchut@igc.irk.ru, lpetrov@igc.irk.ru

Поступила в редакцию 19 марта 2002 г.

На примере стандартных образцов (СО) природных сред сопоставлены алгоритмы установления их метрологических характеристик, регламентируемые различными редакциями ГОСТ 8.532. Показано, что алгоритм редакции 2002 г. снижает влияние резко отличающихся от основной массы результатов и повышает возможность аттестации СО на большее число элементов. Однако при статистической обработке выборок резко асимметричного вида процедура, изложенная в ГОСТ 8.532-2002, не всегда дает достоверную оценку содержания элементов. Тогда необходима экспертиза обрабатываемых данных для выявления недостоверных результатов и исключения их из дальнейших вычислений.

Анчутина Елена Анатольевна – младший научный сотрудник Института геохимии СО РАН.

Область интересов: стандартные образцы состава, межлабораторный эксперимент, статистический анализ.

Автор 12 работ.

Петров Лев Львович – доктор химических наук, академик Метрологической академии, заведующий аналитическим отделом Института геохимии СО РАН.

Область интересов: стандартные образцы состава, аналитические методы и их метрологическое обеспечение, геохимия.

Автор более 180 работ.

Целью аттестации стандартных образцов (СО) является получение метрологических оценок, характеризующих материал СО. Этот процесс регламентируется с 1 марта 2003 г. ГОСТ 8.532-2002, выпущенным взамен ранее действовавшего ГОСТ 8.532-85. Новый стандарт [1] предлагает принципиально иную схему обработки данных, полученных в межлабораторном эксперименте (МЛЭ). Она основана на оценке критического отклонения результатов от медианы выборки данных. В случае, если эта величина больше всех значений абсолютных отклонений результатов измерений от медианы выборки, то за аттестованное значение принимают среднее арифметическое ряда полученных значений концентраций, в ином случае принимают среднее взвешенное ряда, которое вычисляют по формуле

$$\hat{A} = (1/W) \cdot \sum_{i=1}^N w_i \cdot X_{(i)} \quad (1)$$

где W – сумма весовых коэффициентов w_i . В свою очередь, для вычисления w_i находят нормированное отклонение U_i каждого результата $X_{(i)}$ от медианы

$$U_i = d0_i / (5,2 \cdot \text{MAD}0) \quad (2)$$

где $d0_i$ – абсолютные отклонения результатов от медианы и $\text{MAD}0$ – медиана абсолютных ненулевых отклонений. Если $U_i \geq 1$, то w_i приравнивают нулю, в других случаях

$$\omega_i = (1 - U_i^2)^2 \quad (3)$$

Характеристику погрешности межлабораторной аттестации Δ_A оценивают по формуле:

$$\Delta_A = t_{\alpha} \cdot S_A / \sqrt{f+1}, \quad (4)$$

где t_{α} – квантиль распределения Стьюдента со степенями свободы $f = N - 1$ при вычислении по п.5.4. ГОСТ 8.532-2002 (N – количество результатов в выборке) и $f = K - 1$ (K – количество отличных от нуля весовых коэффициентов ω_i) при вычислении по п.5.5. того же ГОСТа [1]; S_A – среднее квадратическое отклонение, вычисляемое по формуле

$$S_A = 1,48 \cdot \text{MAD1}, \quad (5)$$

где MAD1 – медиана абсолютных ненулевых отклонений (в одном случае от среднего арифметического, в другом – от среднего взвешенного).

Согласно ГОСТ 8.532-85 [2] в случае нормального закона распределения полученных в МЛЭ результатов в качестве значения аттестованной характеристики принимали среднее арифметическое ряда, характеристику погрешности межлабораторной аттестации вычисляли по формуле

$$\Delta_{Aco} = \frac{t_{(n-1)} \cdot S}{\sqrt{n}}, \quad (6)$$

где S – стандартное отклонение; $t_{(n-1)}$ – коэффициент Стьюдента; n – число результатов.

Для симметричного (но ненормального) распределения данных или при числе результатов менее 15 аттестованным значением являлась медиана упорядоченного ряда всех возможных полусумм, полученных из результатов выборки; оценкой погрешности межлабораторной аттестации – полуширина доверительного интервала для данной медианы.

Для несимметричного распределения результатов за аттестованную характеристику принимали медиану ряда полученных результатов, в качестве характеристики погрешности – полуширину доверительного интервала для данной медианы.

Как видно, механизмы вычисления аттестуемых характеристик, предлагаемых двумя редакциями стандарта, отличны друг от друга, что может влиять на значение искомых оценок. Чтобы выявить эти различия, была осуществлена обработка полученных для разрабатываемого СО состава гранат-биотитового плагиогнейса ГБПг-1 результатов в соответствии с процедурами статистической обработки, предложенными рассматриваемыми стандартами. Сравнение проводили как по аттестуемым значениям, так и по характеристикам погрешности. Необходимо отметить, что характеристики, полученные по ГОСТ 8.532-85, вычисляли после предварительного анализа выборок данных, что осуществляли согласно подходу, используемому в ИГХ СО РАН [3, 4]. Так как новый стандарт предлагает принципиально иной алгоритм обработки полученных в МЛЭ данных, вычисления согласно процедуре ГОСТ 8.532-2002 проводили по полной выборке.

Аттестованные значения, полученные в результате вычисления с помощью разных алгоритмов статистической обработки результатов МЛЭ (рис. 1), близки или совпадают для выборок данных, имеющих более 20 результатов, и различаются для некоторых микроэлементов, у которых число результатов в выборке менее 20, но более 10.

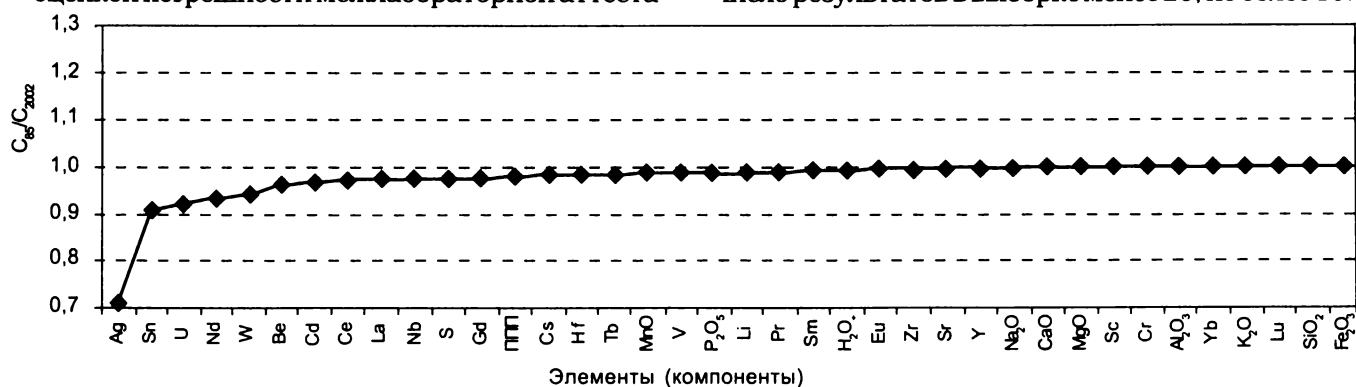


Рис.1. Отношение значений концентраций элементов, полученных при вычислении аттестуемых значений в соответствии с ГОСТ 8.532-85 и ГОСТ 8.532-2002

Максимальные различия наблюдаются для серебра: $C_{\text{ГОСТ 8.532-85}} = 0,17 \text{ г/т}$; $C_{\text{ГОСТ 8.532-2002}} = 0,24 \text{ г/т}$ и германия: $C_{\text{ГОСТ 8.532-85}} = 1,50 \text{ г/т}$; $C_{\text{ГОСТ 8.532-2002}} = 1,19 \text{ г/т}$. Для выборки данных по определению содержания серебра в ГБПг-1 (табл. 1, а) имеется 15 результатов, разброс между минимальным и мак-

симальным значениями составляет два порядка ($C_{\text{min}} = 0,036 \text{ г/т}$; $C_{\text{max}} = 6 \text{ г/т}$). При вычислении согласно ГОСТ 8.532-85 из первоначальной выборки были исключены, как сомнительные, четыре максимальных результата (0,8; 0,8; 2 и 6 г/т), что и дало в итоге установленное значение в два раза

меньшее, чем при вычислении согласно процедуре ГОСТ 8.532-2002. Выборка результатов анализа ГБПг-1 по германию (табл. 1, б) состоит из 13 значений, максимальный и минимальный результат различаются больше, чем в 25 раз (0,6 и 16,1 г/т).

При вычислении согласно ГОСТ 8.532-85 из выборки был удален максимальный результат, равный 16,1 г/т. Распределение симметрично и число результатов менее 15, поэтому за искомое значение принимали медиану всех возможных полусумм ряда, тогда как при вычислении согласно ГОСТ 8.532-2002 – среднее взвешенное ряда (1,19 г/т), которое оказалось близко к медиане (1,21 г/т).

Определенной тенденции к получению более низких или высоких оценок погрешностей аттестации, установленных сравниваемыми алгоритмами обработки результатов, не наблюдается (рис.2, а). Почти для всех макрокомпонентов (кроме P_2O_5 и ППП) более высокие оценки получены согласно процедуре статистической обработки, изложенной в ГОСТ 8.532-85 (рис.2, б). Для ряда элементов, например, для U, Mo и Cd, эта процедура дала более низкие значения погрешностей, в основном, за счет признания части результа-

тов в выборке недостоверными и исключения их перед статистической обработкой (рис.2, в).

Таблица 1
Данные МЛЭ по определению содержания серебра (а) и германия (б) в СО ГБПг-1

а)		б)	
№	C_{Ag} , г/т	№	C_{Ge} , г/т
1	0,0358	1	0,6
2	0,04	2	0,78
3	0,045	3	0,86
4	0,05	4	1
5	0,07	5	1,09
6	0,09	6	1,09
7	0,1	7	1,21
8	0,3	8	1,5
9	0,3	9	2
10	0,3	10	3,1
11	0,4	11	3,6
12	0,8	12	4
13	0,8	13	16,1
14	2		
15	6		

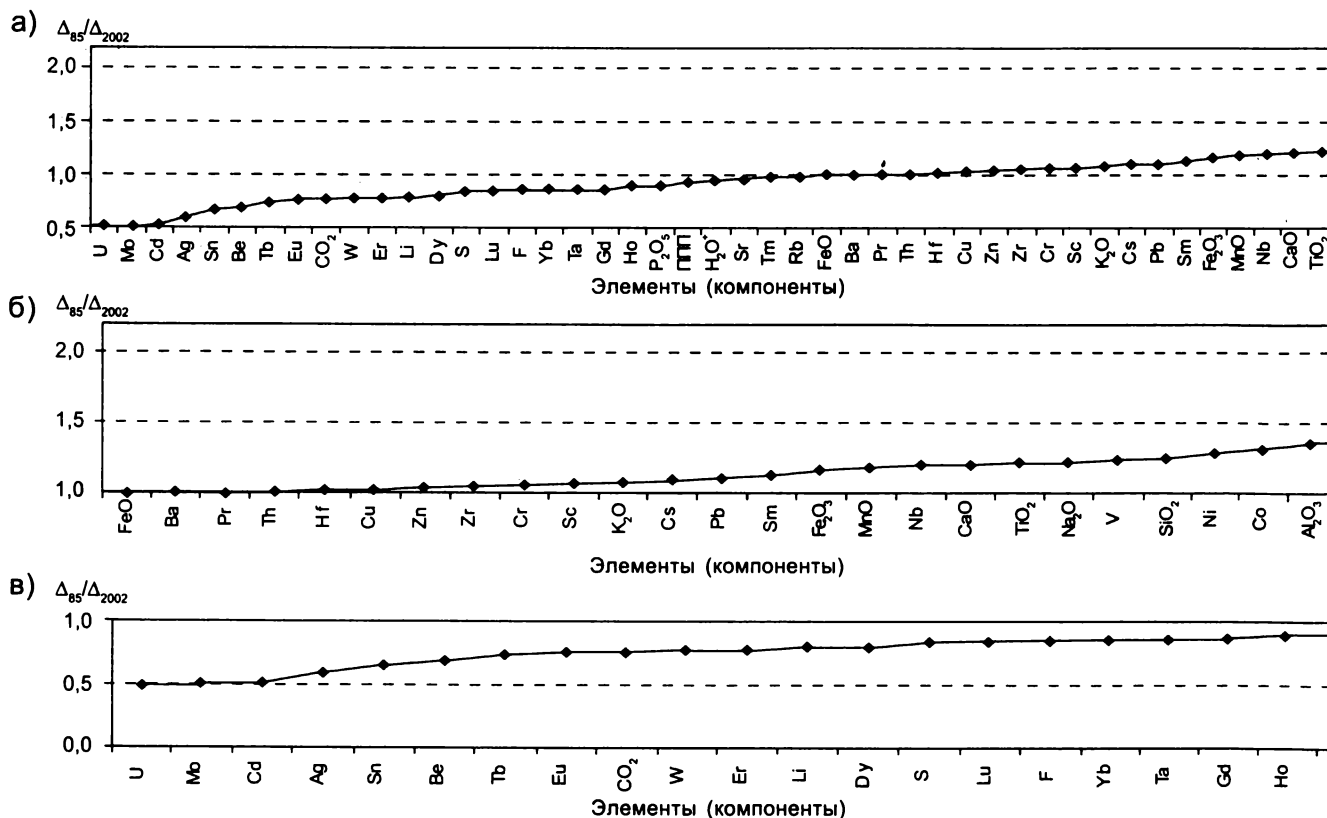


Рис.2. Отношение значений погрешностей аттестации, полученных при вычислении в соответствии с ГОСТ 8.532-85 и ГОСТ 8.532-2002

Если сравнить значения погрешностей аттестованного значения, полученные разными алгоритмами для полной выборки (без отсе-

ния недостоверных с профессиональной точки зрения результатов), новый стандарт работает эффективнее предыдущего, поскольку дает бо-

более низкие значения погрешностей. Это демонстрирует рис.3, на котором, в отличие от данных рис.2, а, погрешности аттестации рассчиты-

таны без предварительного исключения ряда результатов.

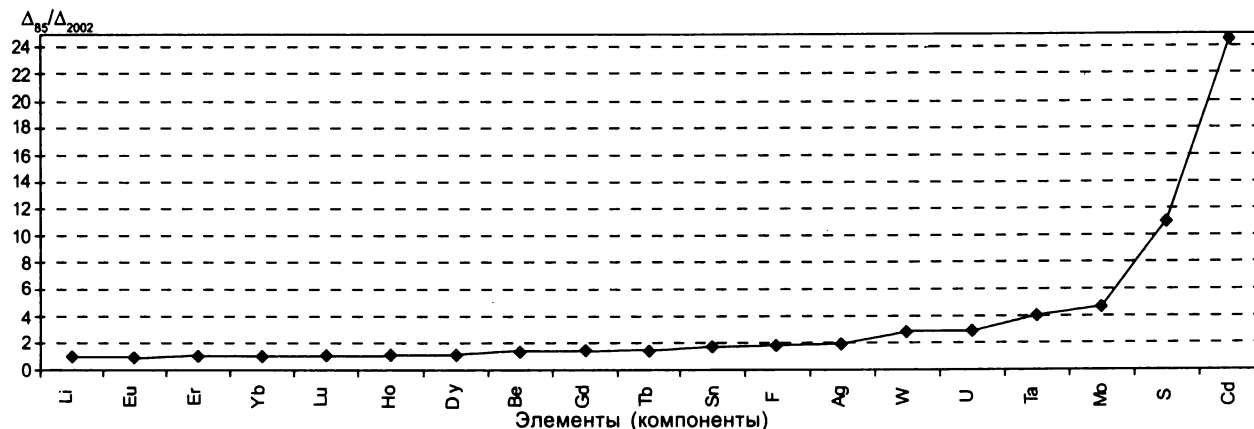


Рис.3. Отношение значений погрешностей аттестуемого значения, полученных при расчетах в соответствии с ГОСТ 8.532-85 и ГОСТ 8.532-2002 для полных выборок данных

Для подавляющего большинства элементов меньшие значения погрешностей дает процедура ГОСТ 8.532-2002, что, несомненно, повышает возможность аттестации СО на большее число элементов. Явное преимущество у нового документа и при обработке выборок, имеющих количество результатов, меньшее 20. Например, после статистической обработки данных МЛЭ по правилам ГОСТ 8.532-85 СО ГБПг-1 не мог быть аттестован по содержанию таллия, для которого имеется 11 результатов, поскольку для правильной с точки зрения разработчиков СО оценки его

содержания необходимо было удалить из выборки два максимальных результата. В таком случае число результатов становится равным девяти, что не соответствует требованиям стандарта к нахождению аттестованных значений (диаграмма распределения результатов представлена рис.4). В тоже время, при вычислении метрологических характеристик по ГОСТ 8.532-2002 погрешность аттестуемого значения равна 0,055 г/т при допустимой 0,064 г/т, что дает возможность аттестовать СО ГБПг-1 по содержанию таллия в ранге государственного стандартного образца.

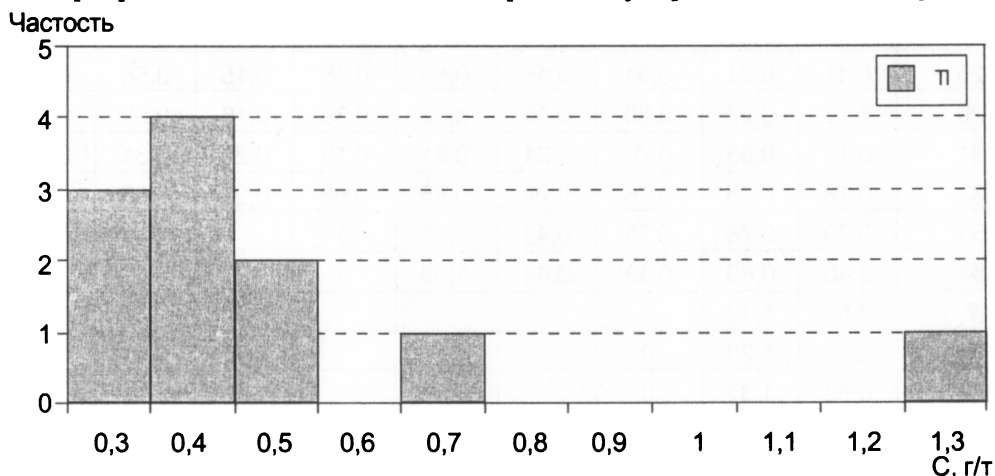


Рис.4. Гистограмма распределения результатов определения содержания таллия в СО ГБПг-1

Наряду с отмеченными выше достоинствами предложенного в ГОСТ 8.532-2002 алгоритма обработки данных МЛЭ, отмечены случаи, когда без предварительной экспертизы выборки данных с целью признания части результатов недостоверными невозможно достигнуть достоверных оценок содержания элемента в аттестуемом образце. Такая картина наблюдается обычно при обработке выборок резко асимметричного и многомодального вида. При этом появление резкой

асимметрии в распределении данных нельзя отнести только за счет плохой организации МЛЭ. Это вызвано реальным уровнем развития аналитики при определении микро- и трудно определяемых элементов (недавно подобная ситуация была описана в [5]).

Рассмотрим, например, статистическую обработку результатов МЛЭ по определению содержания серебра для ГСО состава черного сланца СЧС-1 (разработка ИГХ СО РАН).

Таблица 2

Результаты МЛЭ по серебру для ГСО СЧС-1

№	С, г/т	I			II			III			IV
		d0 _i	U _i	w _i	d0 _i	U _i	w _i	d0 _i	U _i	w _i	d0 _i
1	0,007	0,14	0,28	0,85	0,093	0,36	0,76	0,093	0,36	0,76	
2	0,017	0,13	0,26	0,87	0,083	0,32	0,81	0,083	0,32	0,81	
3	0,02	0,13	0,25	0,88	0,08	0,31	0,82	0,080	0,31	0,82	
4	0,05	0,10	0,19	0,93	0,05	0,19	0,93	0,050	0,19	0,93	
5	0,05	0,10	0,19	0,93	0,05	0,19	0,93	0,050	0,19	0,93	
6	0,075	0,07	0,14	0,96	0,025	0,10	0,98	0,025	0,10	0,98	0,025
7	0,08	0,07	0,13	0,97	0,02	0,08	0,99	0,020	0,08	0,99	0,020
8	0,08	0,07	0,13	0,97	0,02	0,08	0,99	0,020	0,08	0,99	0,020
9	0,09	0,06	0,11	0,98	0,01	0,04	1,00	0,010	0,04	1,00	0,010
10	0,095	0,05	0,10	0,98	0,005	0,02	1,00	0,005	0,02	1,00	0,005
11	0,098	0,05	0,10	0,98	0,002	0,01	1,00	0,002	0,01	1,00	0,002
12	0,1	0,05	0,09	0,98	0	0	1	0	0	1	0
13	0,1	0,05	0,09	0,98	0	0	1	0	0	1	0
14	0,1	0,05	0,09	0,98	0	0	1	0	0	1	0
15	0,1	0,05	0,09	0,98	0	0	1	0	0	1	0
16	0,11	0,04	0,07	0,99	0,01	0,04	1,00	0,01	0,04	1,00	0,010
17	0,11	0,04	0,07	0,99	0,01	0,04	1,00	0,01	0,04	1,00	0,010
18	0,12	0,03	0,05	0,99	0,02	0,08	0,99	0,02	0,08	0,99	0,020
19	0,14	0,01	0,01	1,00	0,04	0,15	0,95	0,04	0,15	0,95	0,040
20	0,15	0	0,01	1,00	0,05	0,19	0,93	0,05	0,19	0,93	
21	0,17	0,03	0,05	0,99	0,07	0,27	0,86	0,07	0,27	0,86	
22	0,17	0,03	0,05	0,99	0,07	0,27	0,86	0,07	0,27	0,86	
23	0,19	0,05	0,09	0,98	0,09	0,35	0,77	0,09	0,35	0,77	
24	0,25	0,11	0,21	0,91	0,15	0,58	0,45	0,15	0,58	0,45	
25	0,26	0,12	0,23	0,89	0,16	0,62	0,39	0,16	0,62	0,39	
26	0,31	0,17	0,33	0,79	0,21	0,81	0,12	0,21	0,81	0,12	
27	0,32	0,18	0,35	0,76	0,22	0,85	0,08	0,22	0,85	0,08	
28	0,53	0,39	0,78	0,15	0,43	1,65	0				
29	0,54	0,40	0,80	0,13	0,44	1,69	0				
30	0,7	0,56	1,12	0							
31	0,74	0,60	1,20	0							
32	0,82	0,68	1,37	0							
33	0,89	0,75	1,51	0							
34	1,3	1,16	2,34	0							
35	1,3	1,16	2,34	0							
36	3,8	3,66	7,40	0							
37	4,2	4,06	8,21	0							
38	5,1	4,96	10,03	0							
М, г/т		0,145			0,100			0,100			0,100
С _к , г/т		0,285			0,15			0,15			0,045
Â, г/т		0,127			0,105			0,105			0,100
Δ _Â , г/т		0,043			0,023			0,017			0,009
Δ _д , г/т		0,025			0,020			0,020			0,019

Примечания: Расчеты выполнены: I - для полной выборки; II - для выборки с №1 по №29; III - для выборки с №1 по №27; IV - для выборки с №6 по №23. № - порядковый номер результата в упорядоченной выборке; С - независимый результат определения содержания серебра в ГСО СЧС-1; $d0_i$ - абсолютные отклонения результатов измерений от медианы; U_i - нормированное отклонение от медианы (по формуле (2)); w_i - весовой коэффициент (по формуле (3)); М - медиана выборки; C_k - критическое отклонение результата от медианы; \hat{A} - установленное аттестованное значение (для I, II, III - среднее взвешенное ряда, для IV - среднее арифметическое); Δ_d - погрешность аттестации (по формуле (4)); Δ_d - допустимая погрешность.

Установление аттестованных значений концентраций благородных металлов в материалах природных сред – всегда нетривиальная задача из-за многомодальности и резкой асимметрии вида распределения получаемых по ним результатов [4, 6]. Разброс в выборке из 38 результатов составляет три порядка (табл. 2). Как было указано выше, при статистической обработке согласно ГОСТ 8.532-2002 в случае, когда хотя бы одно из отклонений результатов от медианы превышает критическое значение (что характерно для большинства асимметричных видов распределения данных), за аттестованное значение \hat{A} принимают среднее взвешенное ряда, вычисляемое по формуле (1). Так как на его значение влияют весовые коэффициенты w_i , которые приравниваются нулю в случае, когда нормированное отклонение результата более 1 (формула (2)), резко отличающиеся от основной массы данных результаты не участвуют в установлении \hat{A} и вносят лишь вклад в погрешность.

В данном случае критическое отклонение C_k (0,285 г/т) меньше наибольшего абсолютного отклонения результатов измерений от медианы $d0_i$ (4,96 г/т), поэтому за аттестованное значение принимают среднее взвешенное ряда, для чего вычисляют нормированное отклонение от медианы U_i по (2). Девять значений $U_i > 1$ (колонка

I), поэтому весовые коэффициенты w_i для них приравниваются нулю, и они не дают вклад в результат вычисления \hat{A} (формула (1)). Полученные оценки не соответствуют требованиям аттестации (погрешность аттестации больше допустимой).

Как видно из гистограммы распределения результатов (рис. 5), эти девять значений находятся в "хвосте" максимальных значений и соответствуют четырем крайним модам. Обнуляя весовые коэффициенты для таких результатов, процедура ГОСТ 8.532-2002 естественным образом исключает их из окончательной выборки или делает их влияние минимальным. Если отсечь их как недостоверные и пересчитать характеристики (колонка II), медиана выборки уменьшится (с 0,145 г/т до 0,100 г/т), и, соответственно, изменятся и другие характеристики. В частности, снова два значения $U_i > 1$ (мода напротив значения 0,55 г/т по оси X, рис. 5), и оценка погрешности Δ_d вновь превышает допустимую Δ_d . Вычисление аттестуемых значений после отсечения и этих двух результатов дополнительно (колонка III) не изменило значения медианы и аттестованной характеристики, но дало значение погрешности аттестации (0,017 г/т) меньше допустимого (0,020 г/т). Все $U_i < 1$, и, таким образом, все результаты дают вклад в значение среднего взвешенного и погрешности.

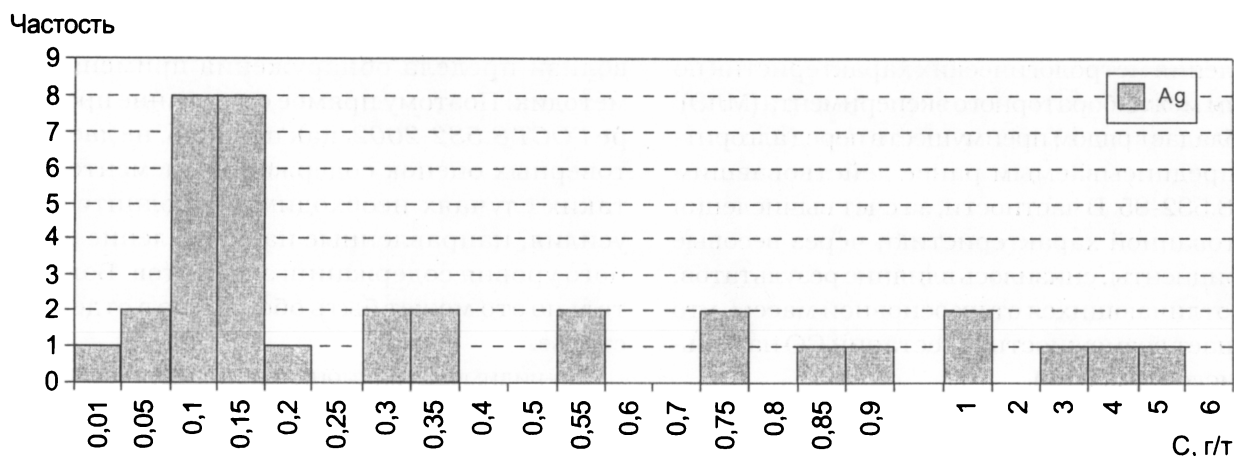


Рис. 5. Гистограмма распределения результатов определения содержания Ag в образце СЧС-1

При определении аттестованного содержания серебра в ГСО СЧС-1 разработчиками были исключены как недостоверные все данные, находящейся вне диапазона данных упорядоченного ряда от 0,075 до 0,14 г/т, и обработку проводили для выборки, состоящей из четырнадцати ре-

зультатов. Это обосновывали тем, что в этом диапазоне находились результаты наиболее авторитетных лабораторий, имеющих опыт в анализе золото- и серебросодержащих проб и обладающих методиками с низкими пределами обнаружения [4], а также то, что здесь сосредоточились резуль-

таты, полученные процедурами, связанными с анализом большой навески (более 10 г). Аттестованной характеристикой, вычисленной согласно ГОСТ 8.532-2002, в данном случае (колонка IV) является среднее арифметическое (так как C_k больше любого d_0), и полученные характеристики соответствуют требованиям аттестации ($\hat{A} = 0,100$ г/т; $\Delta_A = 0,009$ г/т; $\Delta_d = 0,019$ г/т). Аттестованные характеристики, вычисленные согласно ГОСТ 8.532-85, следующие: $\hat{A} = 0,100$ г/т; $\Delta_A = 0,010$ г/т.

Рис. 6 отображает изменение относительных погрешностей с урезанием этой же выборки данных по определению содержания серебра в ГСО СЧС-1, начиная с максимальных значений (с 1-го по 38-ое, с 1-го по 37-ое ... с 1-го по N-ое значение упорядоченного ряда). Пик в виде трех точек (38, 37, 36 по оси X) связан с тем, что до удаления из выборки трех максимальных значений, соответствующих крайней моде на рис.

5, распределение не отвечало нормальному и симметричному, и за аттестованное значение принимали медиану выборки, а за погрешность - полуширину доверительного интервала для данной медианы. Далее с отсечением "правого хвоста" распределения наблюдается постепенное уменьшение относительной погрешности и в районе приблизительно с $N=24$ по $N=16$ ее стабилизация, что для гистограммы (рис.5) соответствует диапазону максимальной моды и достижению симметрии в распределении данных. Увеличение относительной погрешности после $N=16$ вновь говорит о возникновении асимметрии, только на это раз правосторонней. На графике также видно, что в интервале, когда число результатов более 20, относительные погрешности, вычисленные согласно ГОСТ 8.532-2002, неизменно ниже аналогичных оценок, полученных по ГОСТ 8.532-85.

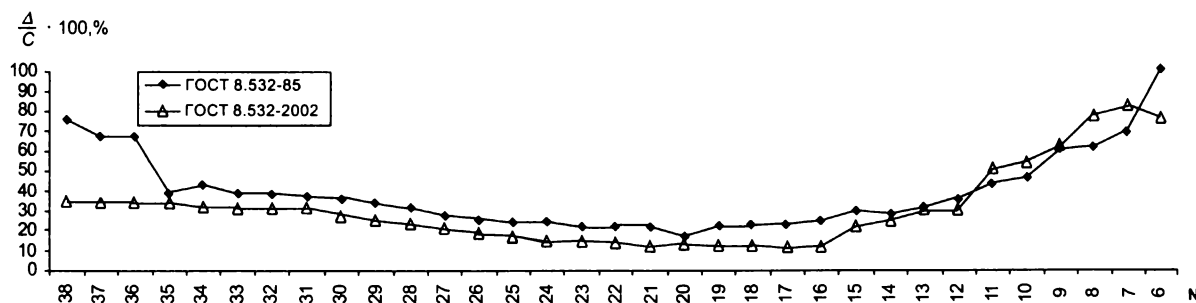


Рис. 6. Изменение относительных погрешностей, вычисленных согласно ГОСТ 8.532-85 и ГОСТ 8.532-2002 для выборок упорядоченных результатов от 1 до N (Ag в ГСО СЧС-1)

Заключение

ГОСТ 8.532-2002 предложил новую схему установления метрологических характеристик по данным межлабораторного эксперимента (МЛЭ). Она обладает рядом преимуществ перед алгоритмом, предписываемым ранее действовавшим ГОСТ 8.532-85. В частности, за счет вычисления аттестованной характеристики через весовые коэффициенты, снизилось влияние результатов, резко отличающихся от их основной массы, что повышает возможность аттестации СО на большее число элементов.

Тем не менее, в ряде случаев, например, при разработке СО состава природных сред, виды распределения полученных данных часто резко асимметричны (или даже многомодальны), главными причинами чего являются сложный мак-

росостав образцов и то, что содержание большинства определяемых микроэлементов находится вблизи предела обнаружения применяемых методик. Поэтому прямое следование процедуре ГОСТ 8.532-2002, как правило, не дает достоверных оценок содержания элементов, и в таких случаях необходимы дополнительные усилия, направленные на осмысление реального уровня содержания элементов. Более детально это может быть обсуждено в отдельной статье.

По-видимому, рационально в следующей редакции ГОСТ 8.532 акцентировать внимание пользователя на процесс предварительной экспертизы выборки данных, чтобы позволить применять подходы, не предписанные напрямую стандартом (конечно, с неопременным их обоснованием).

ЛИТЕРАТУРА

1. ГОСТ 8.532-2002. Стандартные образцы состава веществ и материалов. Межлабораторная метрологическая аттестация. Содержание и порядок

проведения работ. Минск, ИПК Изд-во стандартов, 2002. 9 с.

2. ГОСТ 8.532-85 (СТ СЭВ 4570-84). Стандартные об-

разцы состава веществ и материалов. Порядок межлабораторной аттестации. М: Изд-во стандартов, 1985. 15 с.

3. Петров Л.Л. Закономерности распределения результатов в аналитических интервалах методик выполнения измерений при количественных методах элементного анализа // Заводская лаборатория. 2001. Т.67, № 12. С. 49-58.

4. Petrov L.L. Multielement standard samples of Black Shale/ L.L.Petrov, Yu.N.Kornakov, I.Ya.Korotaeva,

E.A.Anchutina et.c. // Geostandards and Geoanalytical Research. 2004. V.28, N1. P. 89-103.

5. Karpov Yu. A. Interlaboratory comparisons between results from testing metallurgical products for precious metal content/ Yu.A.Karpov, T.A.Makarova, D.F.Makarov and A.P.Kuznetsov // Accreditation and Quality Assurance. 2004. V.9, N7. P.404-410.

6. Juvonen R. Analysis of gold and the platinum group elements in geological samples / Geological Survey of Finland (ESPOO). 229 p.p.

* * * * *

COMPARISON OF DATA TREATMENT ALGORITHMS STATED IN NATIONAL STANDARDS GOST 8.532-85 AND GOST 8.532-2002

E.A.Anchutina, L.L.Petrov

Algorithms of metrological characteristic assignment regulated different versions of the Normative Document, GOST 8.532, are compared by the example of environmental reference materials (RMs). It is shown that the algorithm of GOST 8.532-2002 reduces the influence of results being distinctly different from bulk results and the possibility of RM certification for more elements is increasing. Nevertheless, sometimes direct conformity of calculation algorithm stated in GOST 8.532-2002 does not give reliable estimations, for example, when sets having extreme asymmetric data distribution view are processed. To correctly interpret the data obtained preliminary analysis of data set is necessary to assume some results as unreliable and to eliminate them from the following statistic treatment.